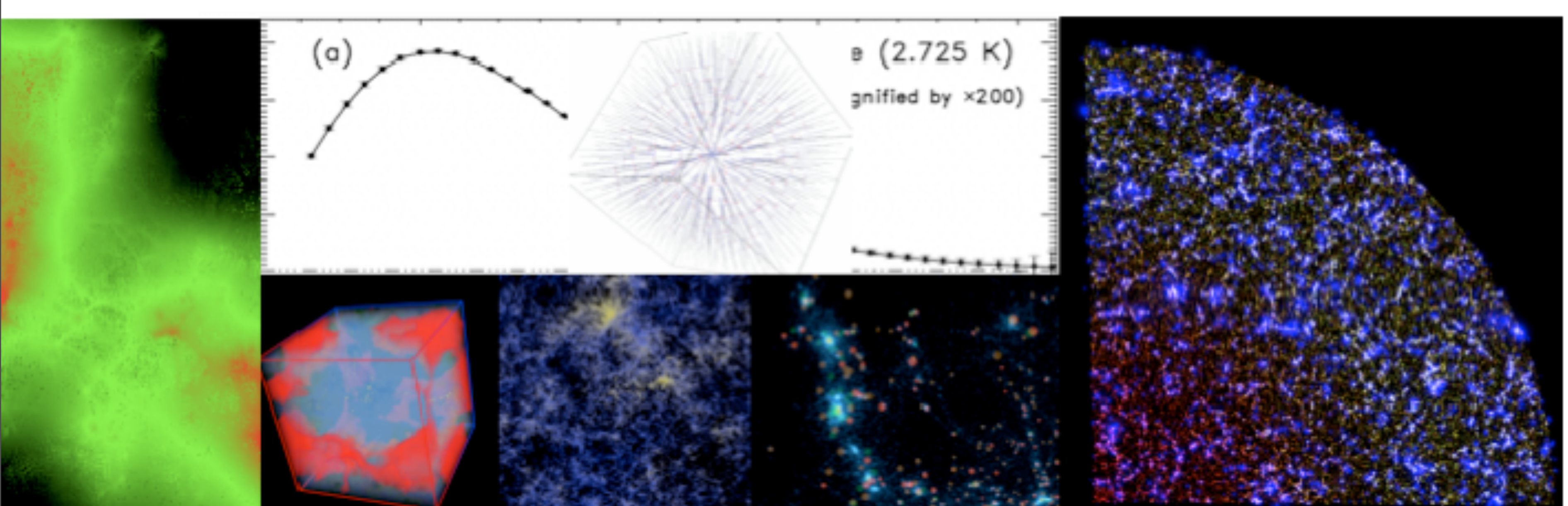
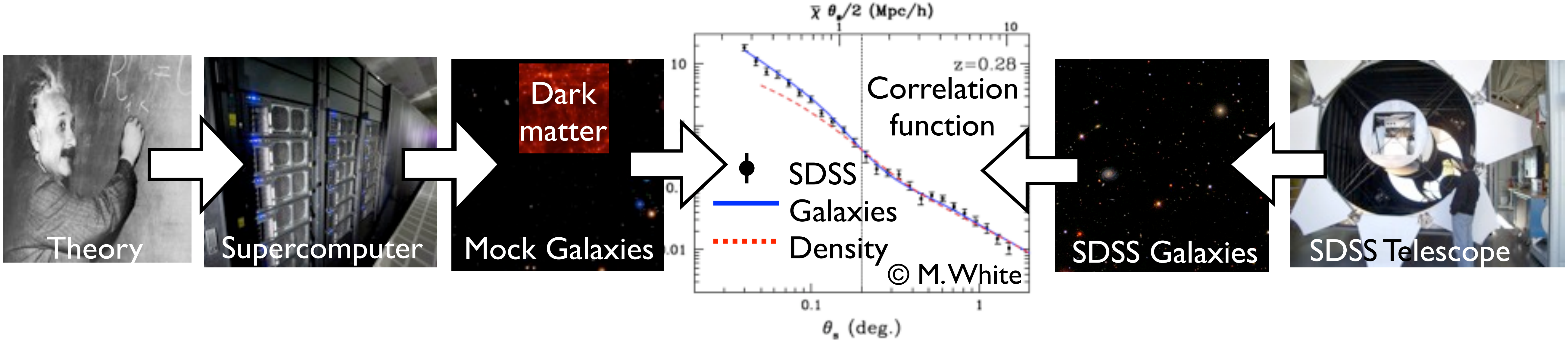
Large Scale Structure Simulations

Katrin Heitmann

Argonne National Laboratory
High Energy Physics/
Mathematics & Computer Science Division



The Role of N-body Simulations in Cosmology

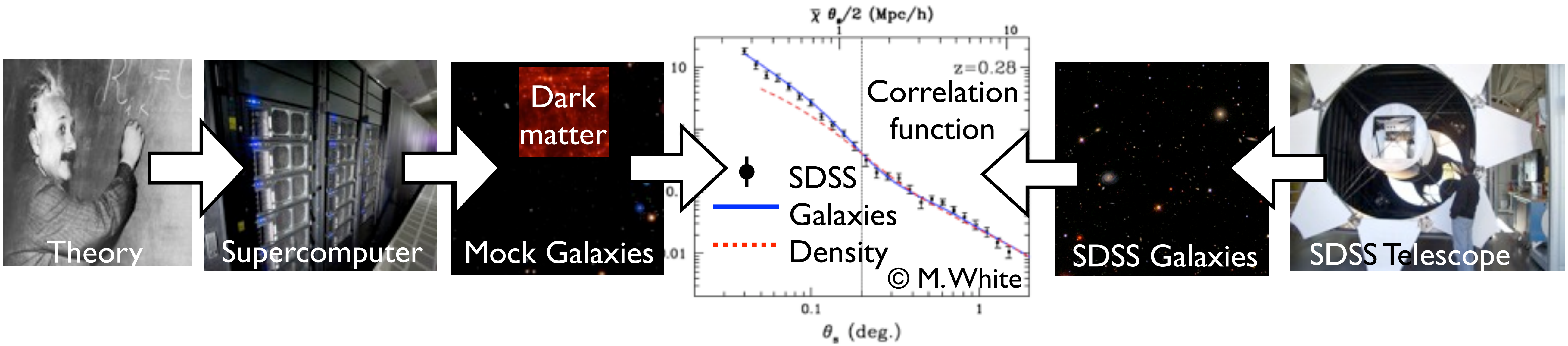


All detailed knowledge of structure formation in the nonlinear regime comes from N-body simulations

- Discovering new physics in the nonlinear regime
- Exploring new cosmological probes
- Analysis of survey results with precision predictions
- Solving cosmological inverse problems
- Basis for generating mock observations
- Estimation of measurement errors

U.S. DEPARTMENT OF ENERGY | Office of Science

The Role of N-body Simulations in Cosmology



All detailed knowledge of structure formation in the nonlinear regime comes from N-body simulations

- **Discovering new physics in the nonlinear regime**
- **Exploring new cosmological probes**
- **Analysis of survey results with precision predictions**
- **Solving cosmological inverse problems**
- Basis for generating mock observations
- Estimation of measurement errors

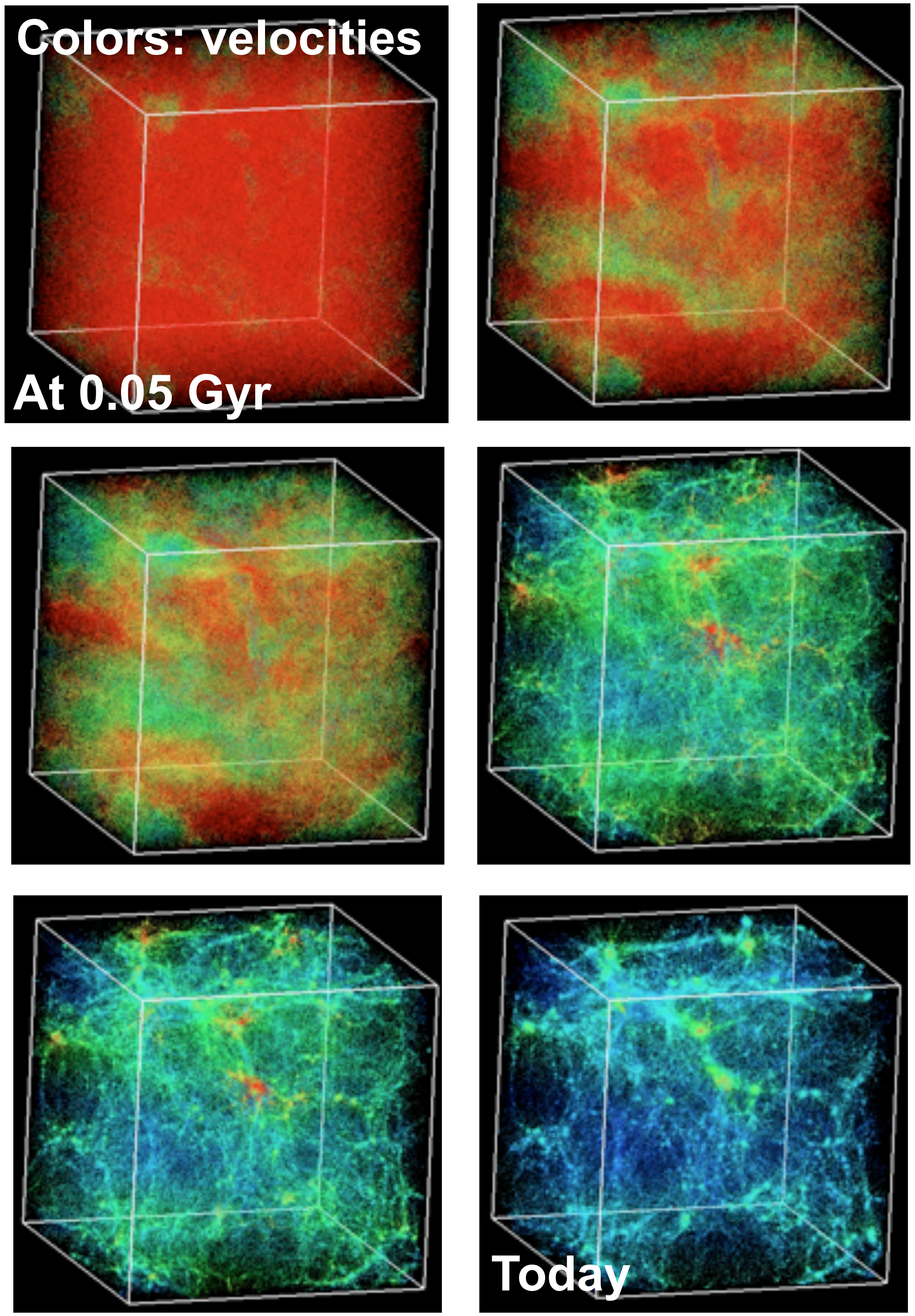
} Risa's talk

U.S. DEPARTMENT OF ENERGY | Office of Science

N-body Simulation Basics

- Follow evolution of dark matter tracer particles in an expanding universe
- Solve Vlasov-Poisson equation
- In principle: N^2 problem (N = number of particles), in practice, approximate methods based on particle and grid solvers
  - Tree methods: Gadget, HOT, PKDGRAV
  - Adaptive mesh refinement: ART, Enzo, FLASH, **Nyx**, Ramses
  - Hybrid: Gadget-2, **TreePM, HACC**

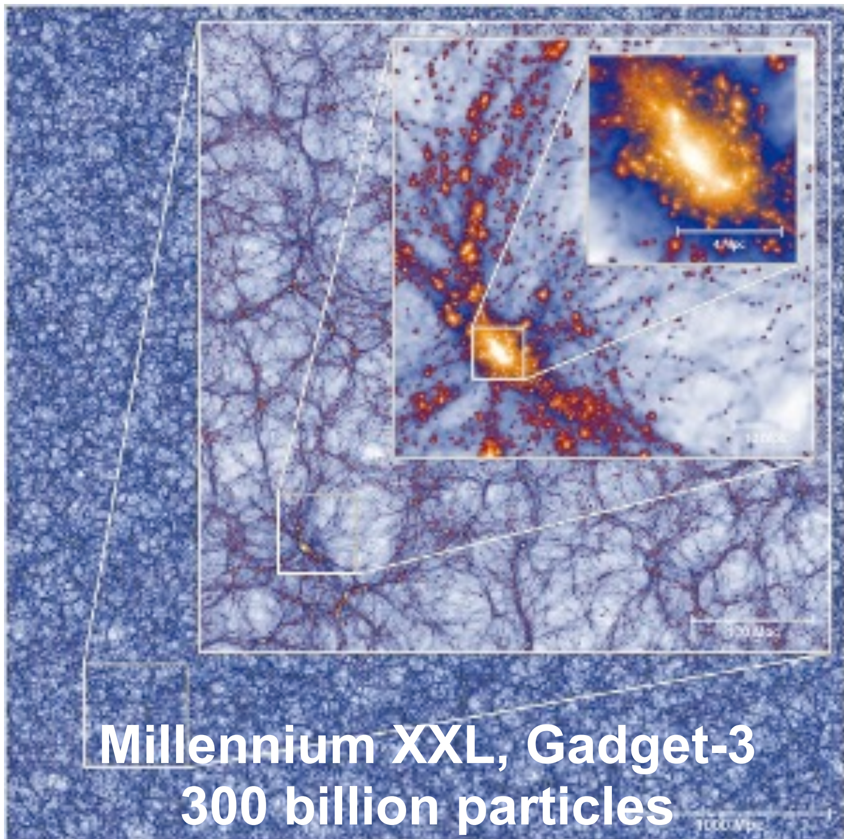
$$
\begin{aligned}
\frac{\partial f_i}{\partial t} + \dot{\mathbf{x}}\frac{\partial f_i}{\partial \mathbf{x}} - \nabla\phi\frac{\partial f_i}{\partial \mathbf{p}} &= 0, \qquad \mathbf{p} = a^2\dot{\mathbf{x}}, \\
\nabla^2\phi &= 4\pi G a^2(\rho(\mathbf{x},t) - \langle\rho_{\rm dm}(t)\rangle) = 4\pi G a^2 \Omega_{\rm dm}\delta_{\rm dm}\rho_{\rm cr}, \\
\delta_{\rm dm}(\mathbf{x},t) &= (\rho_{\rm dm} - \langle\rho_{\rm dm}\rangle)/\langle\rho_{\rm dm}\rangle), \\
\rho_{\rm dm}(\mathbf{x},t) &= a^{-3}\sum_i m_i \int d^3\mathbf{p} f_i(\mathbf{x},\dot{\mathbf{x}},t).
\end{aligned}
$$



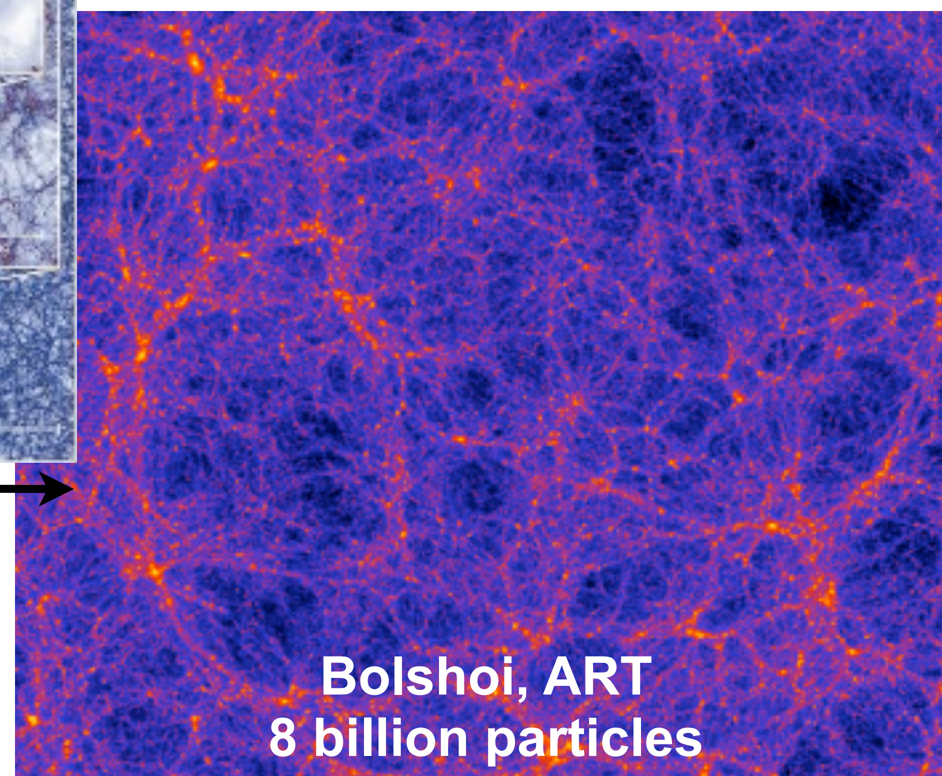
The Virtues of N-body Simulations

- **Advantages:**
  - Capture relevant physics well out into the nonlinear regime
  - Cover large dynamic range: from Gigaparsecs down to less than a kiloparsec, therefore allows coverage of survey size volumes
  - Underlying physics is well understood
  - Several proven numerical paradigms exist, these allow exploration of new high-performance architectures
  - No free physics parameters, code parameters can be set to reach sub-percent accuracy
  - Simulations can be carried out relatively fast (days to weeks time scale), therefore allow exploration of cosmological parameter space, many realizations can be carried out
  - Post-processing via halo occupation distribution models or semi-analytic models allows exploration of additional physics in a straightforward way
- **Disadvantage:**
  - Miss physics on small scales (gas dynamics, feedback..., see Nick's talk), require phenomenological techniques based on observations (self-calibration)

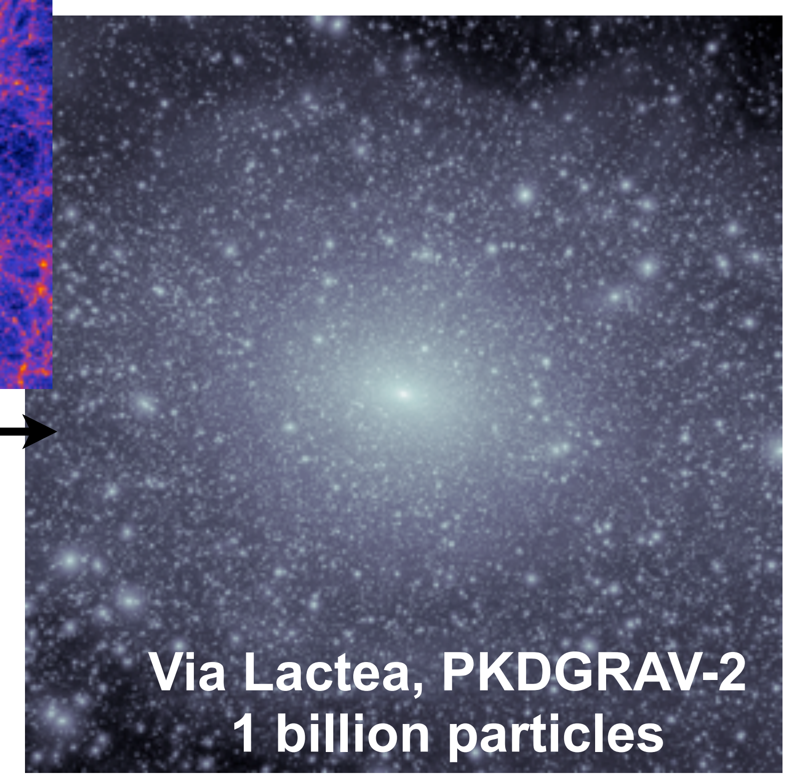
Physics Results on All Scales



4100 Mpc/h



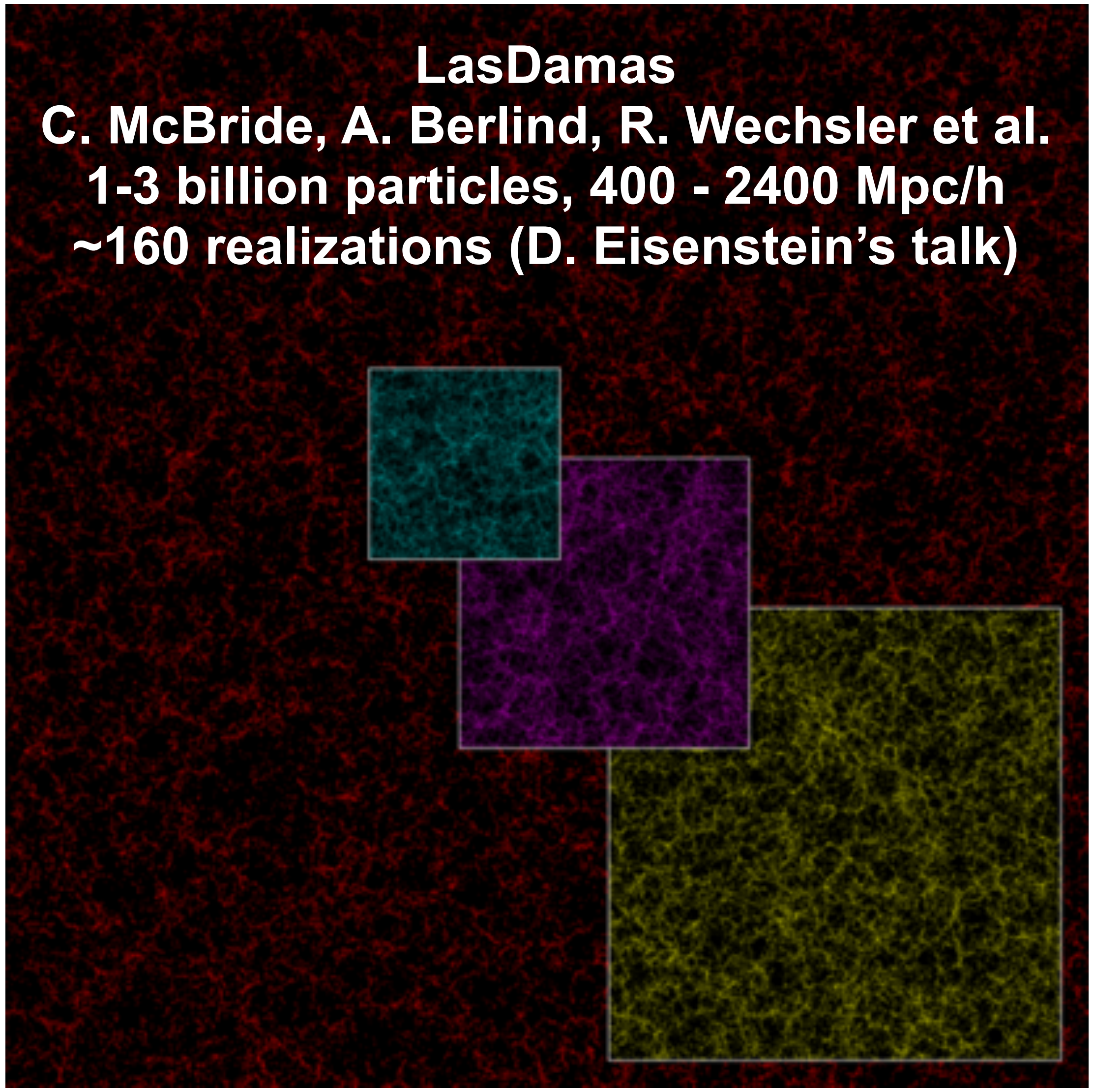
250 Mpc/h



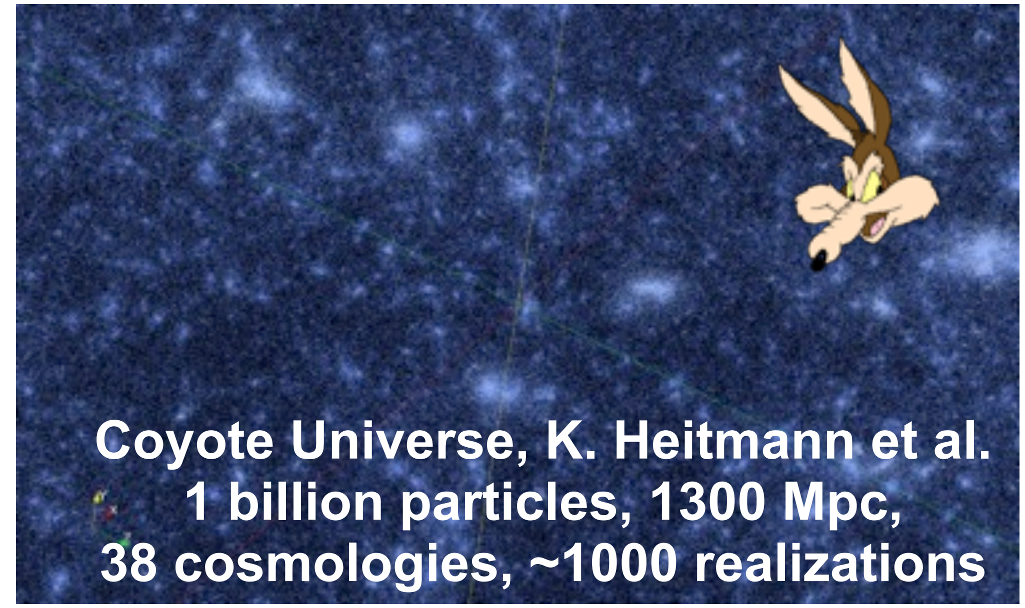
0.8 Mpc/h

U.S. DEPARTMENT OF ENERGY

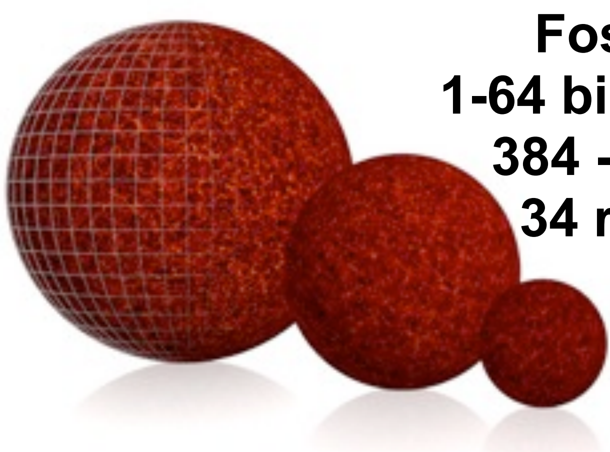
Surveys Need Statistics!



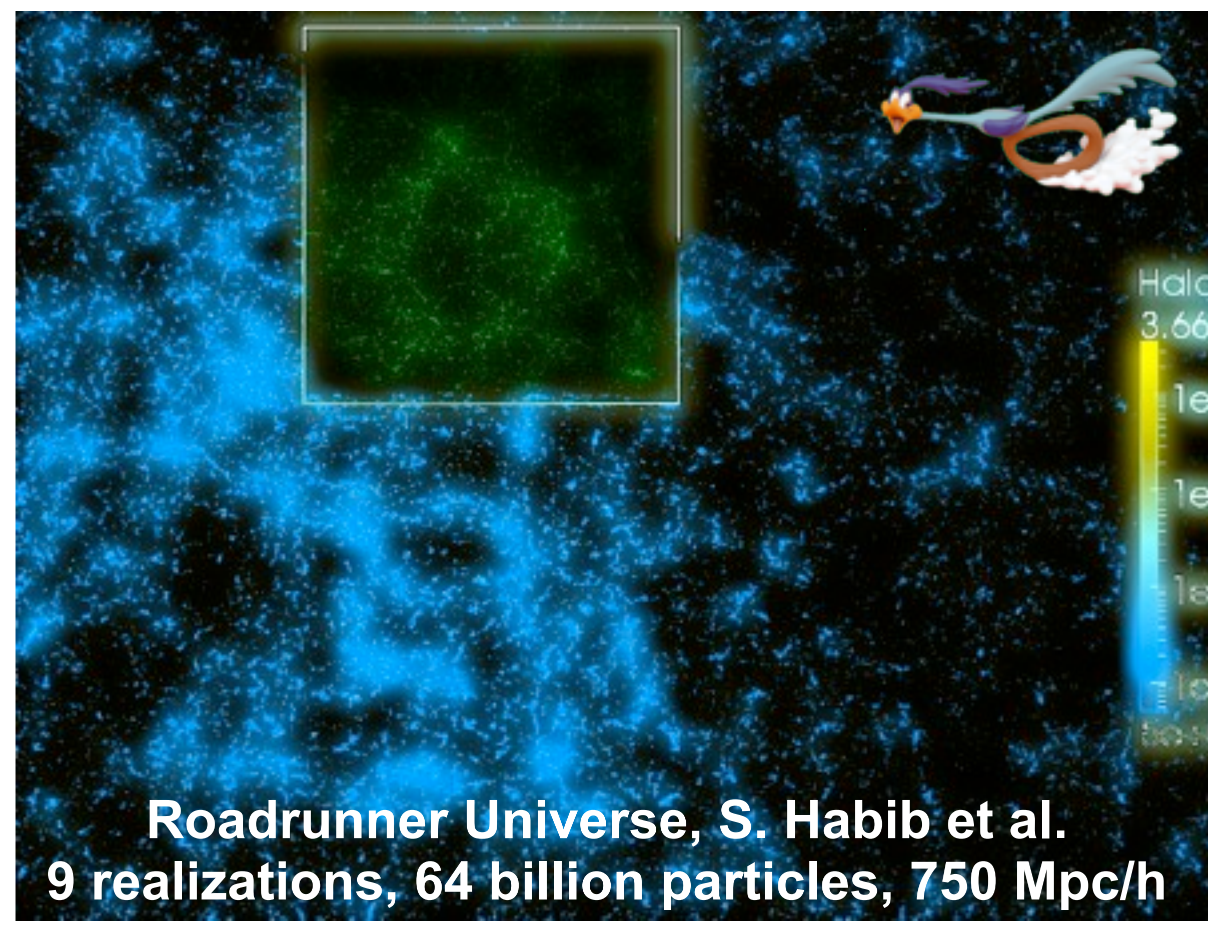
LasDamas
C. McBride, A. Berlind, R. Wechsler et al.
1-3 billion particles, 400 - 2400 Mpc/h
~160 realizations (D. Eisenstein's talk)



Coyote Universe, K. Heitmann et al.
1 billion particles, 1300 Mpc,
38 cosmologies, ~1000 realizations



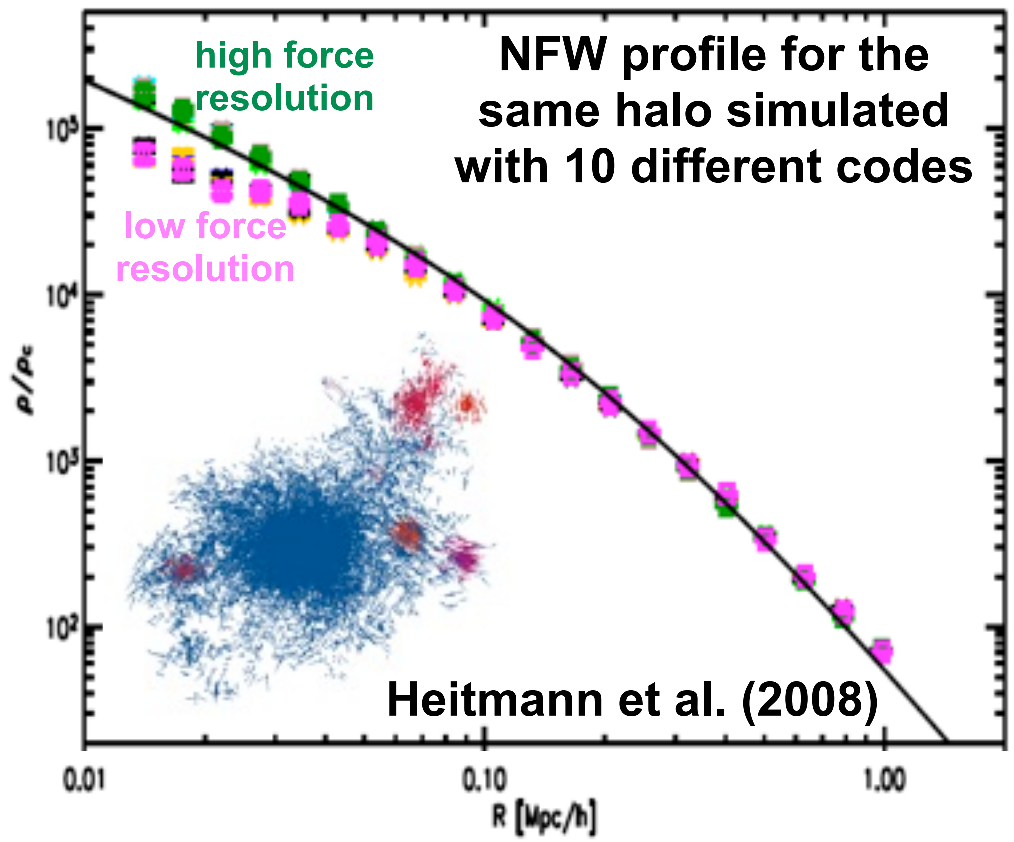
Fosalba et al.
1-64 billion particles,
384 - 7680 Mpc/h
34 realizations
MICE

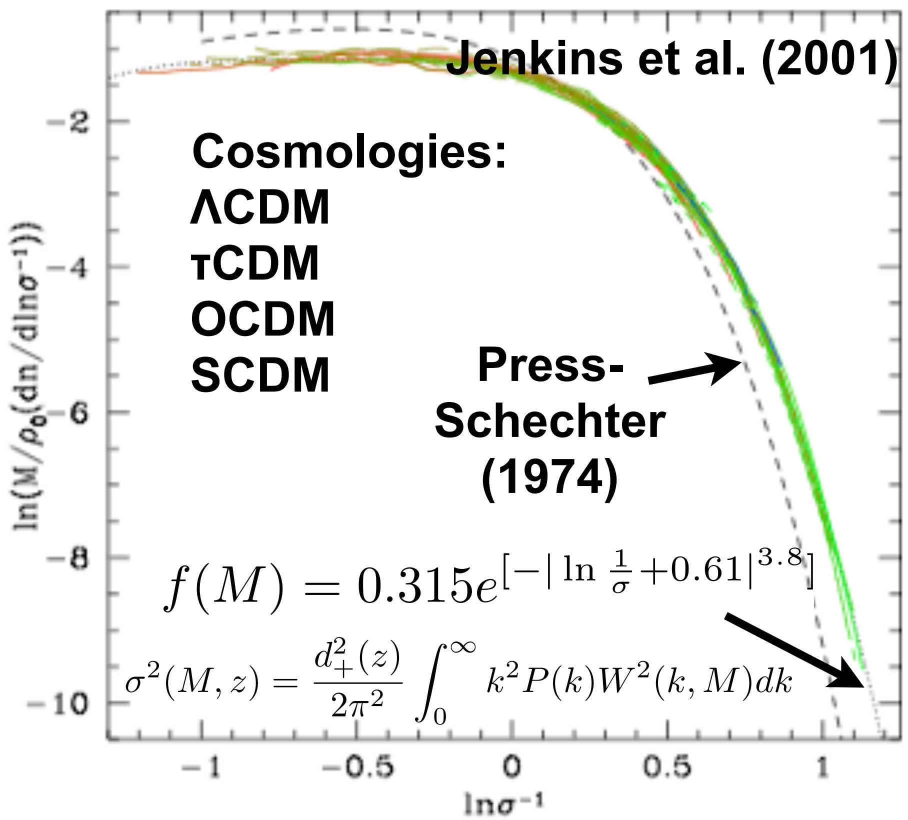


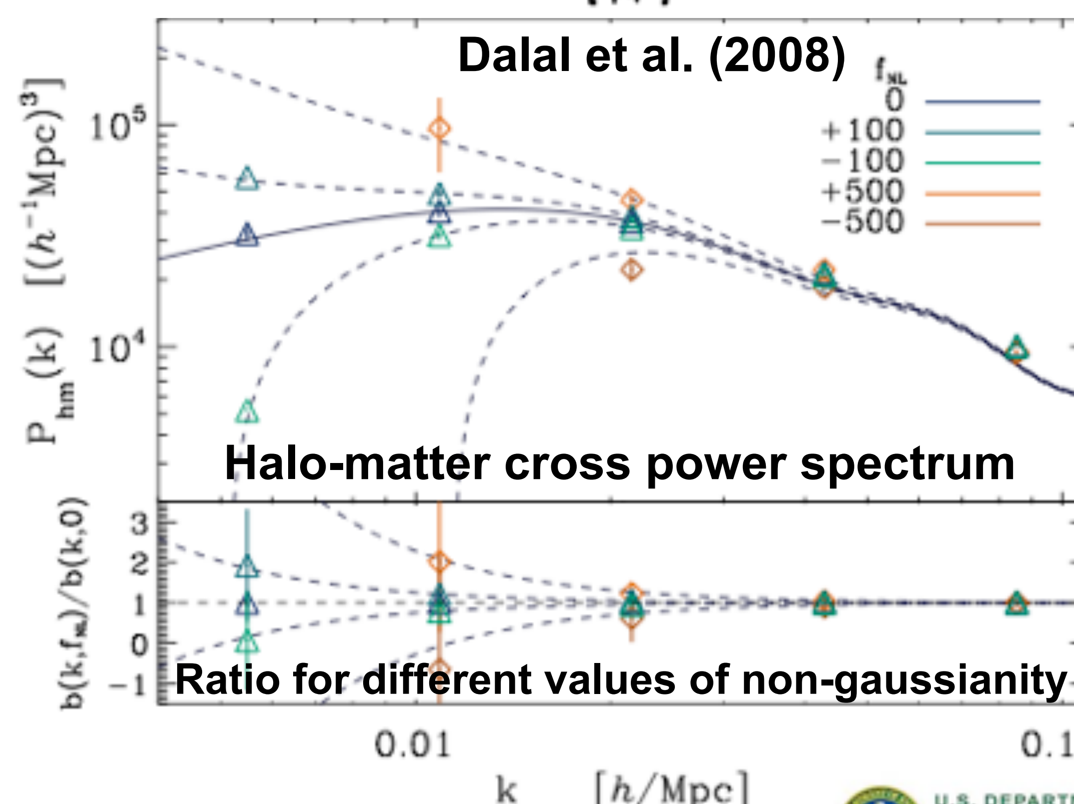
Roadrunner Universe, S. Habib et al.
9 realizations, 64 billion particles, 750 Mpc/h

Discovering New Physics in the Nonlinear Regime

- **Halo profiles:** Navarro, Frenk & White (NFW, 1996, 1997) find that the density profile of a halo as a function of radius has the same shape independent of mass and cosmology
- **Universality of the mass function:** For a particular halo definition (friends-of-friends with a linking length of b=0.2) and expressed via the variance of the linear density field $\sigma^2(z,M)$, mass function is independent of epoch/cosmology at the 20% level
- **Non-gaussianity/bias:** Non-gaussianity effects clustering of dark matter halos, bias is strongly scale dependent

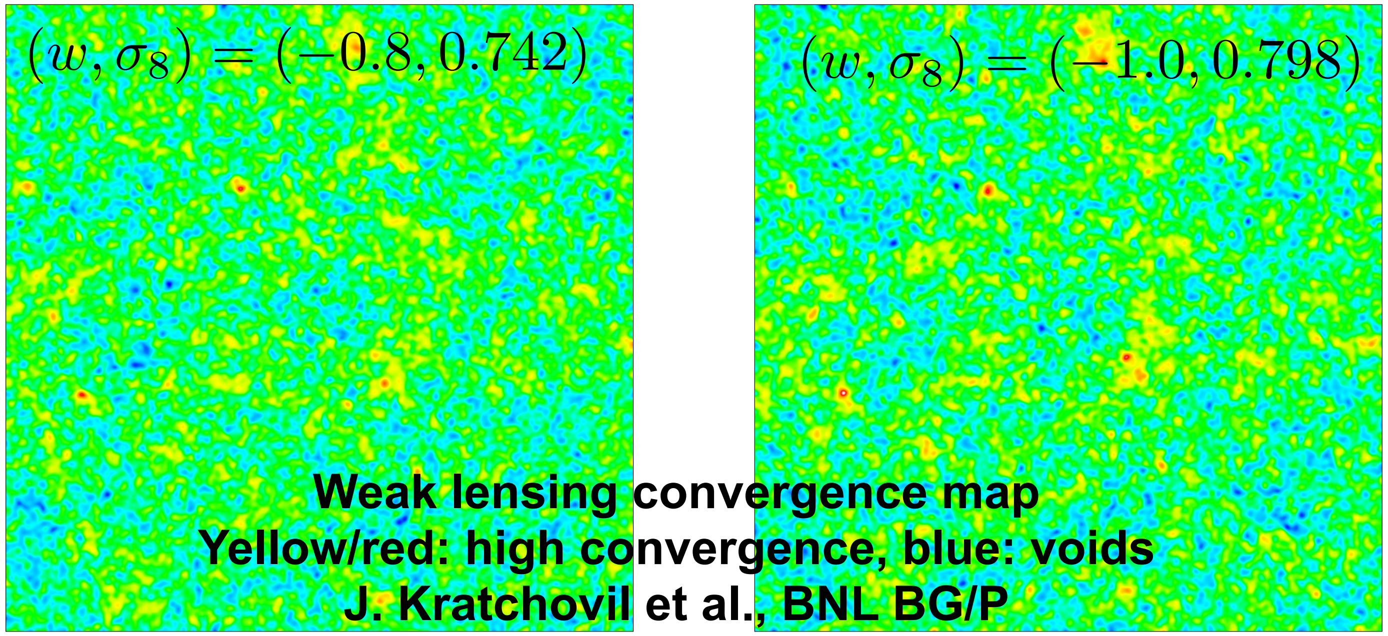




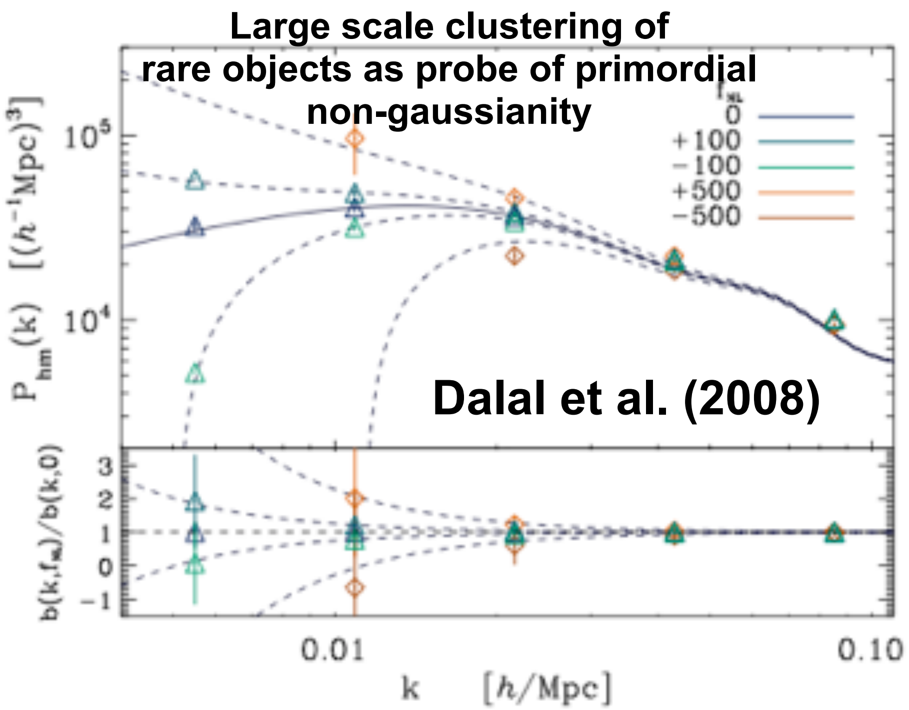


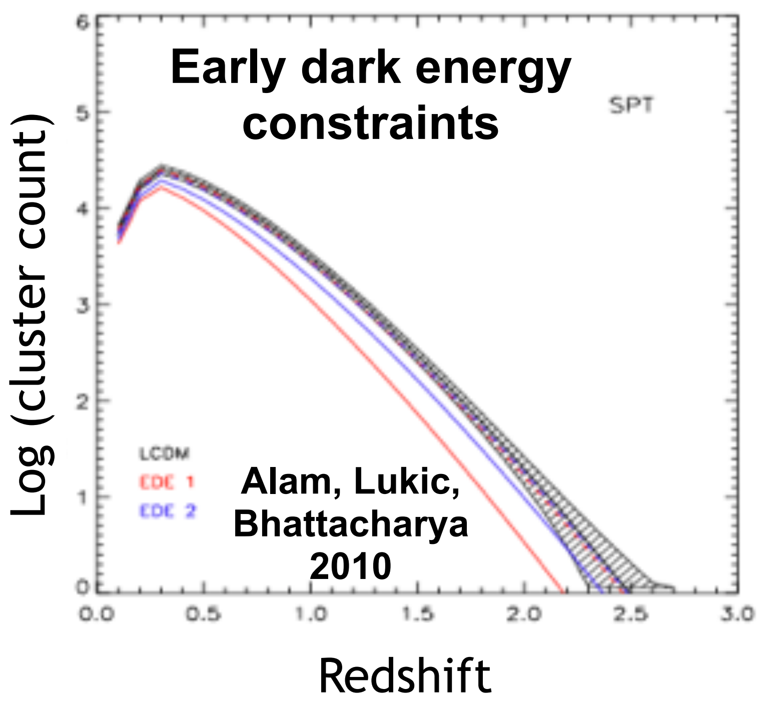
Exploring New (and Established) Cosmological Probes

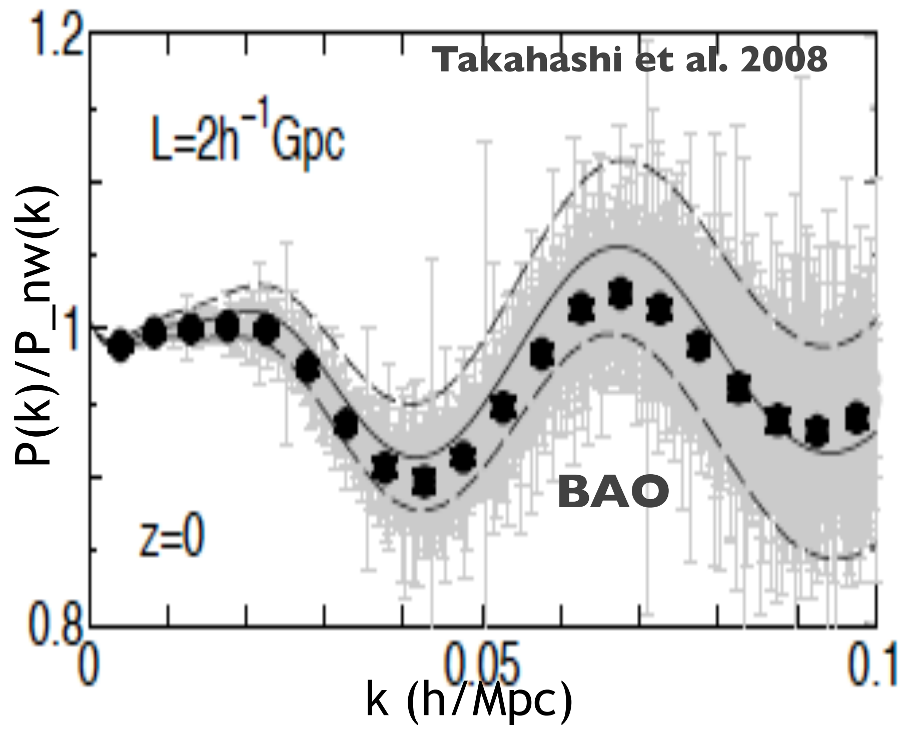
- Sensitivity of cosmological probes to new physics (e.g. early dark energy, modified gravity)
- Estimate of the strength of expected signal over noise
- New cosmological probes



Weak lensing convergence map
Yellow/red: high convergence, blue: voids
J. Kratchovil et al., BNL BG/P



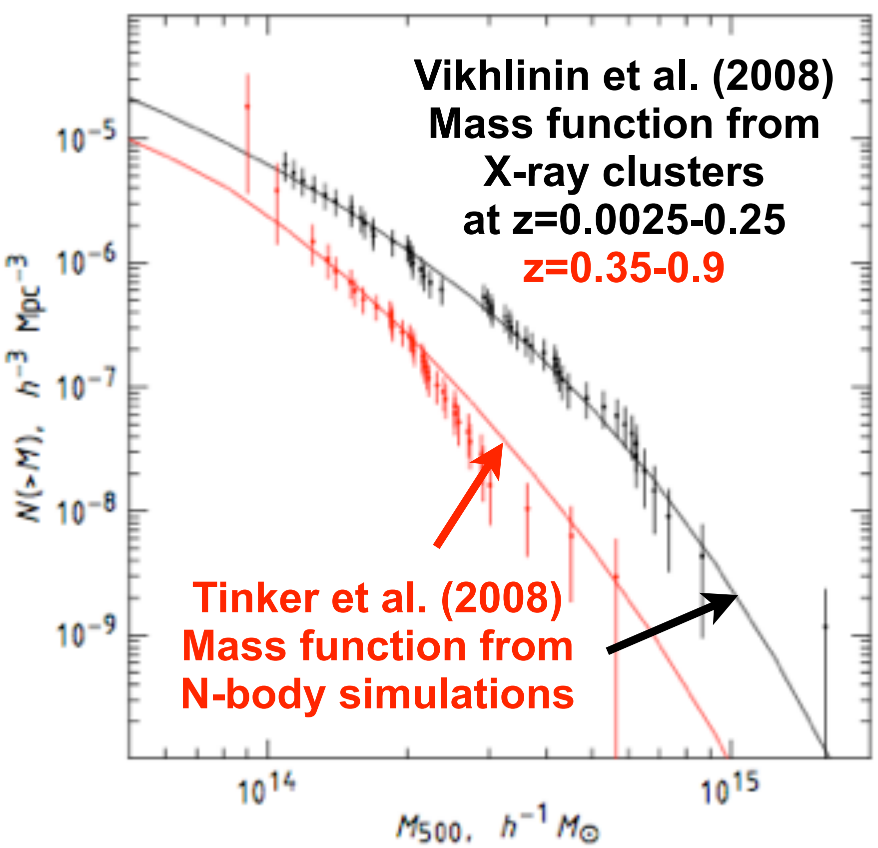


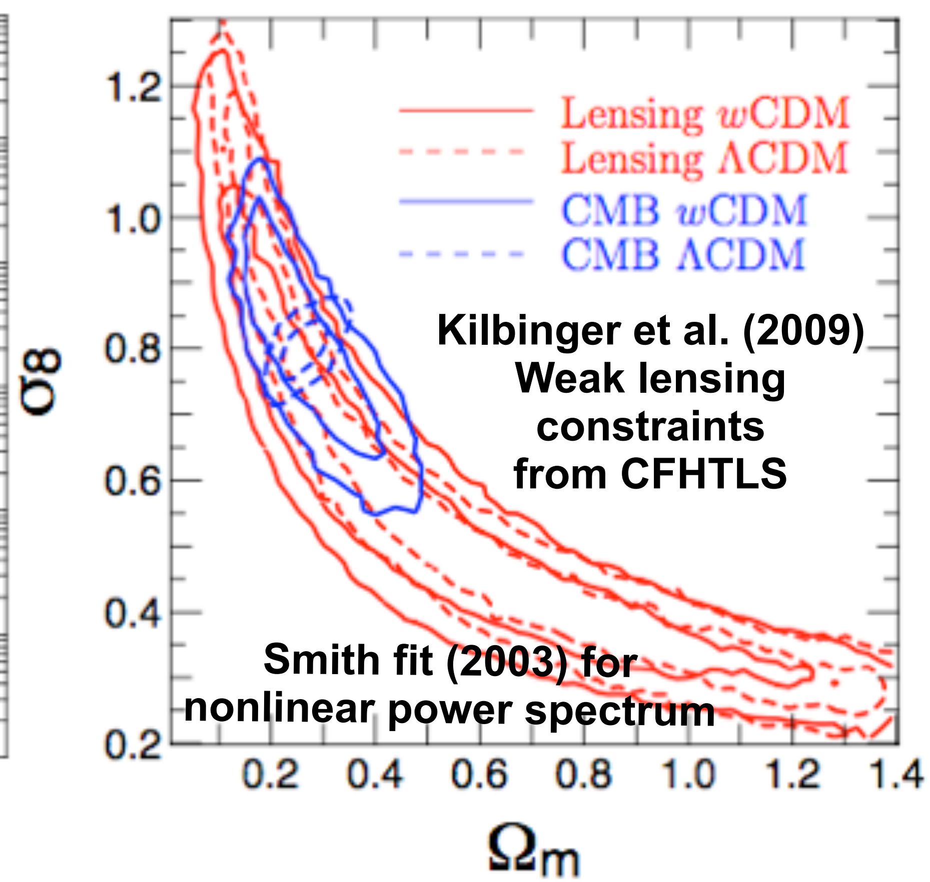


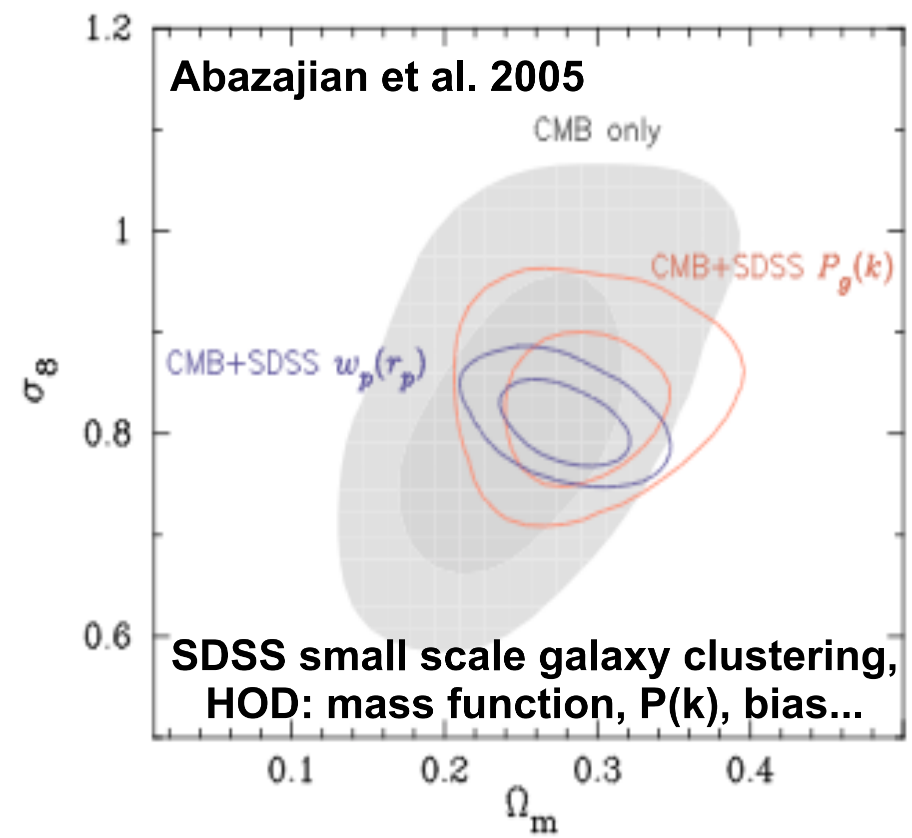
U.S. DEPARTMENT OF ENERGY
Office of Science

Analysis of Survey Results: Three Examples

- **Cluster Cosmology:** Cosmological constraints from Chandra X-ray cluster observations, mass function matched to Tinker et al. 2008 fit from N-body simulations
- **Weak lensing:** Cosmological constraints from Canada-France-Hawaii-Telescope Legacy Survey (CFHTLS), Markov Chain Monte Carlo calculations based on Smith (Halo) fit for nonlinear power spectrum (calibrated with N-body simulations)
- **Small scale galaxy clustering:** Cosmological constraints from Sloan Digital Sky Survey, modeling based on Halo Occupancy Distribution (HOD), includes calibration of mass function, power spectrum, bias from N-body simulations



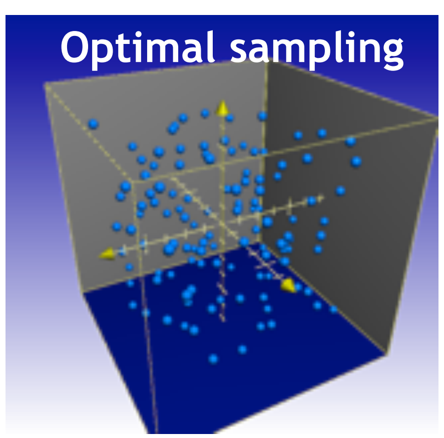


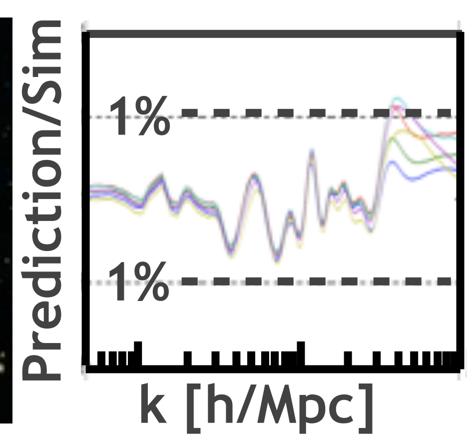


U.S. DEPARTMENT OF ENERGY | Office of Science

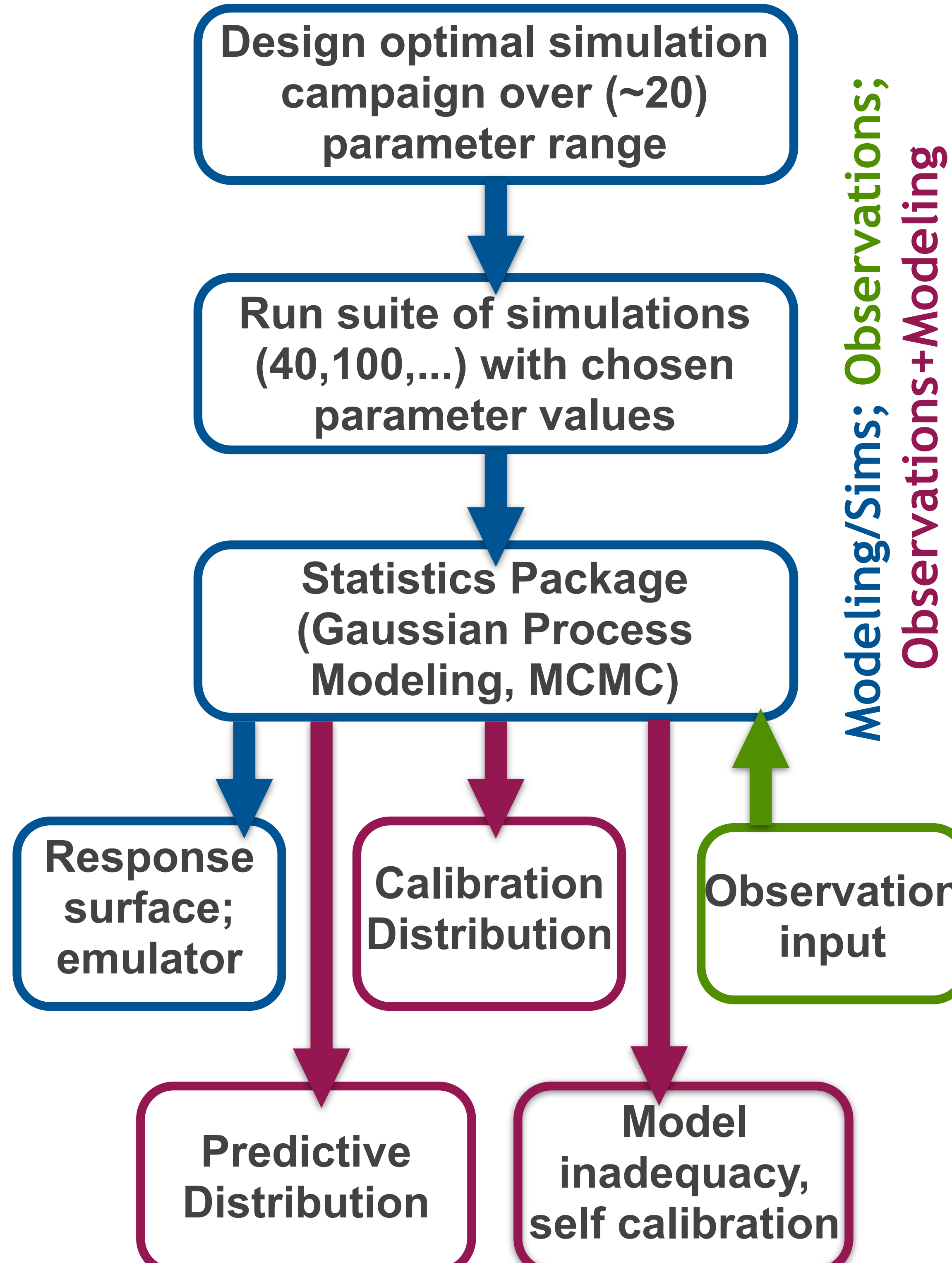
Solving the Inverse Problem

- **Challenge:** To extract cosmological constraints from observations in non-linear regime, need to run Marko Chain Monte Carlo code; input: 10,000 - 100,000 different models
- **Current strategy:** Fitting functions for e.g. P(k), accurate at 10% level, not good enough!
- **Brute force:** Simulations, ~30 years on 2000 processor cluster...
- **Only alternative:** emulators





Heitmann et al. 2006, Habib et al. 2007

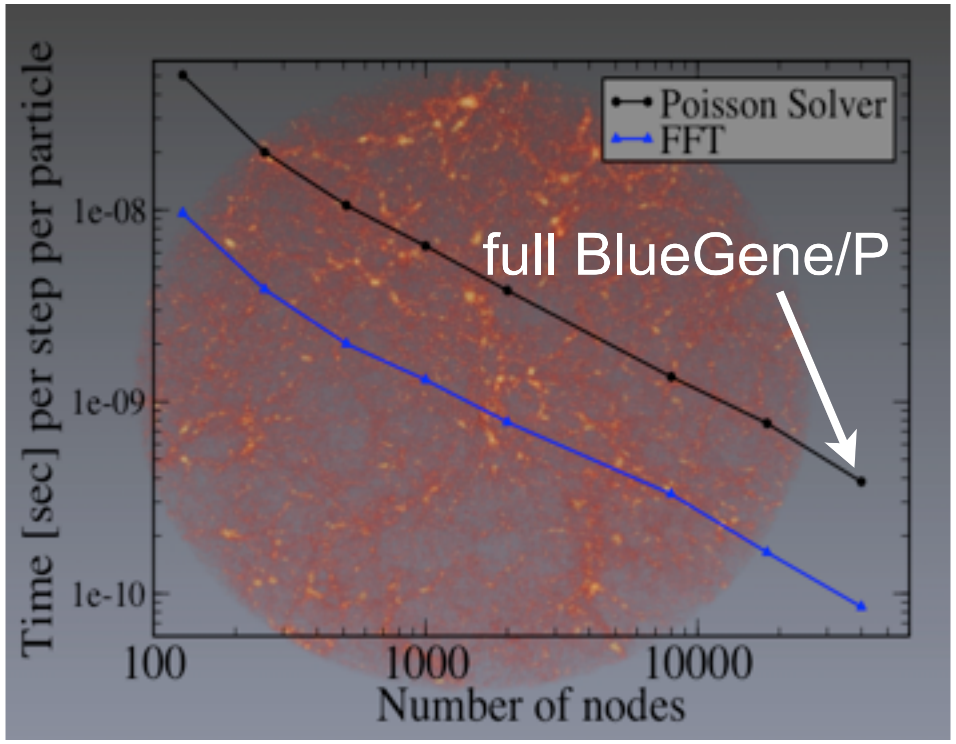


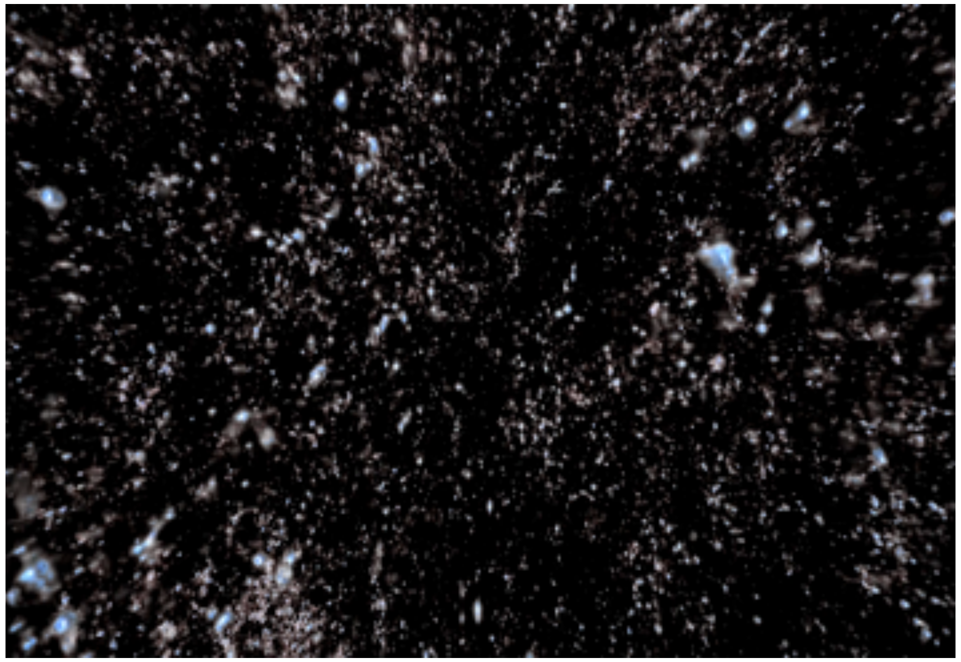
Cosmic Emulator in Action: LSSFast

- Instantaneous ‘oracle’ for nonlinear power spectrum, reduces compute time from weeks to negligible, accurate at 1% out to k~1/Mpc for wCDM cosmologies
- Enables direct MCMC with results from full simulations for the first time

Hybrid Simulation Codes: HACC and TreePM

- **H**ardware **A**ccelerated **C**osmology **C**ode (HACC) (Habib et al. 2009, Pope et al. 2010)
  - New cosmology code designed with major aim of easy portability between different high-performance computing architectures
  - Long-range force: particle mesh, short-range force: particle-particle or tree, architecture dependent
  - Strong collaboration between Los Alamos and Argonne National Labs in code development
  - First version ran on Roadrunner, Cell-accelerated machine, second version on GPU cluster (OpenCL implementation); at SC10: run on heterogenous system with in-situ visualization, now: port to BlueGene system
  - First science: BAO in the Lyman-alpha forest (9 PM runs, 64 billion particles each), precision cosmology with high-resolution code (16 billion particles and up)
- **TreePM** (White 2002)
  - Well established N-body code, used for very large number of science projects: weak lensing, galaxy clustering, BAO, mock catalogs, ...

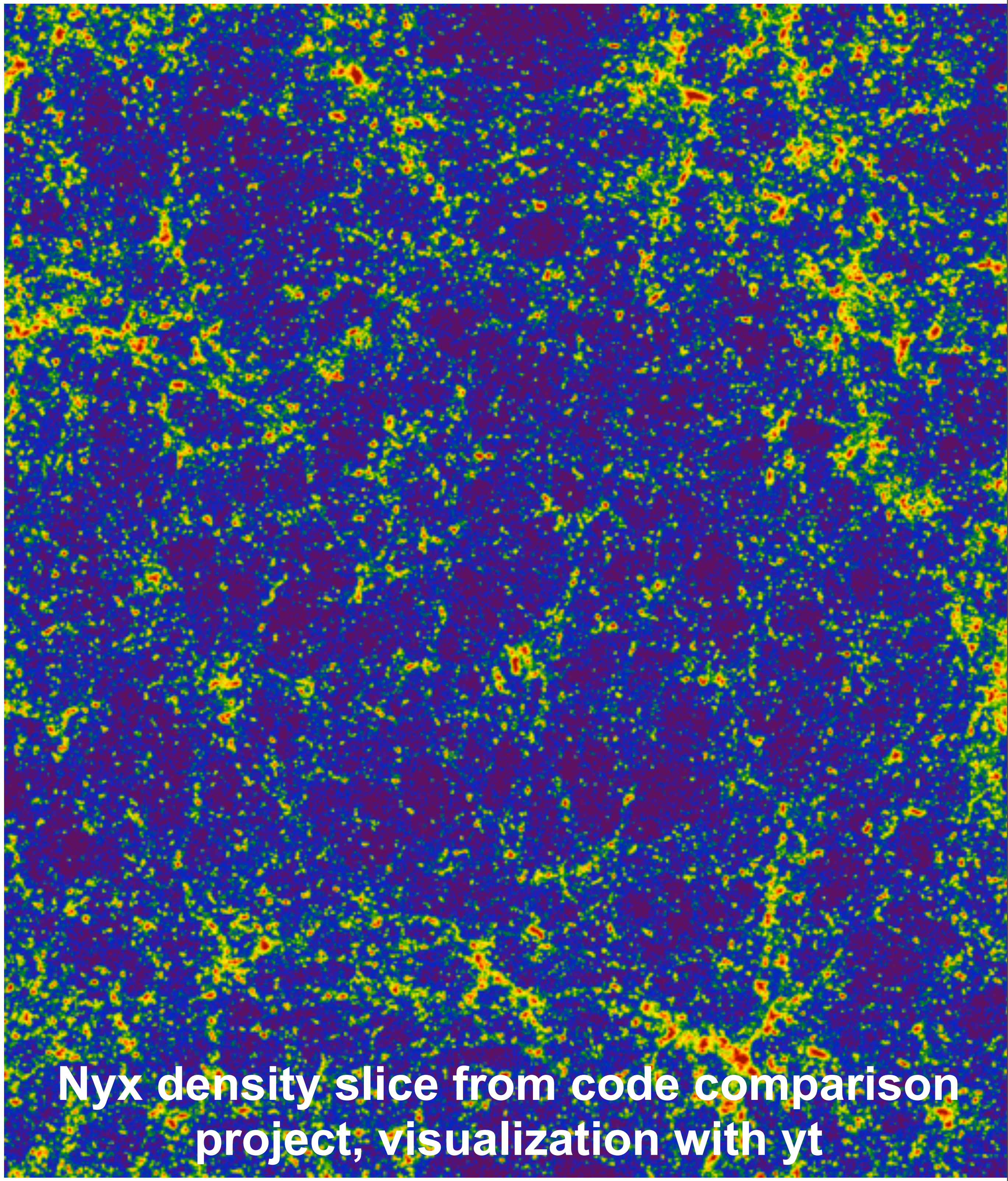




Halos from Roadrunner simulation, visualization: J. Woodring

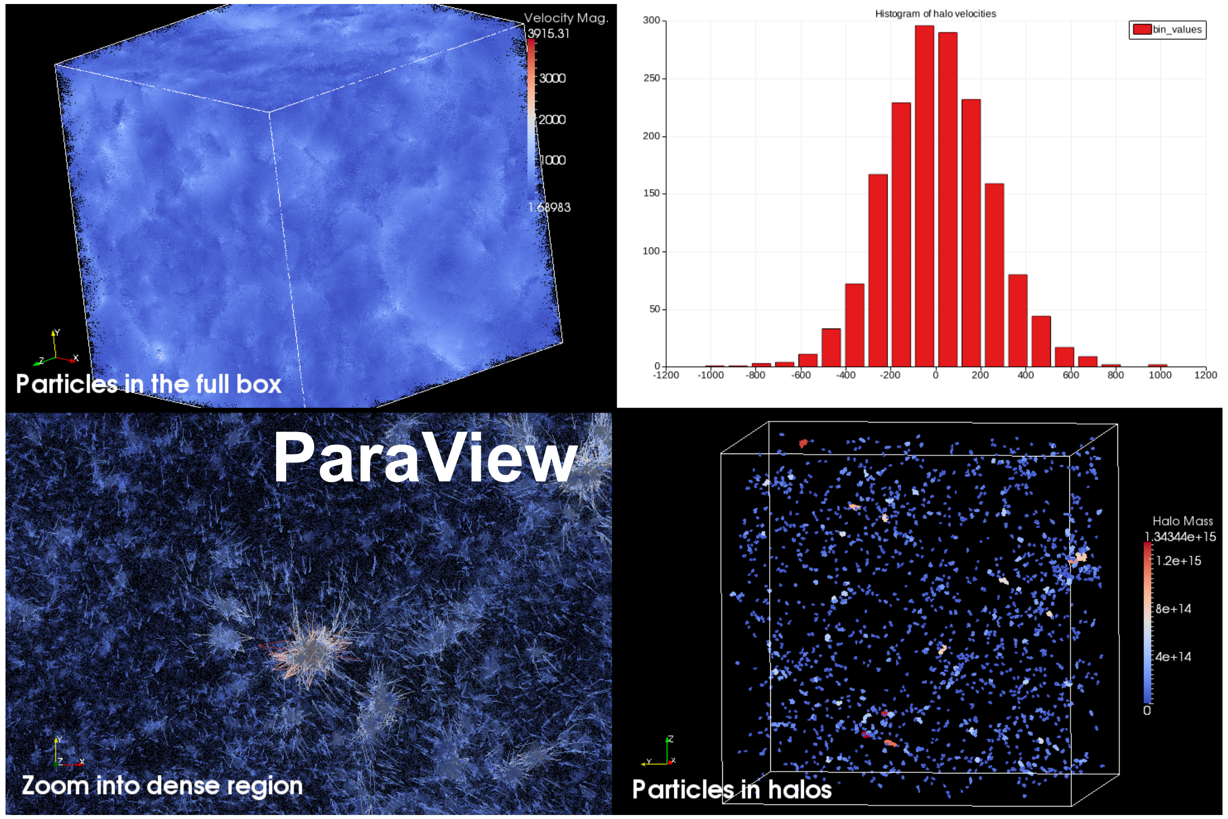
Adaptive Mesh Simulation Codes

- **Nyx: New adaptive mesh refinement (AMR) code developed at LBNL**
- **Block-structured AMR with multigrid Poisson solver**
- **Main focus for AMR codes (also Enzo, ART): hydrodynamics simulations, discussed in detail in Nick's talk**



Analysis Capabilities: ParaView and yt

- **Analysis of simulations as demanding as carrying out the simulations**
- **Visualization-aided analysis**
  - Visualization is important for understanding the results, and for verification & validation
  - Visualization and analysis have to go hand-in-hand to ensure efficiency
- **ParaView (open source)**
  - Based on VTK (Visualization Tool Kit), cosmological analysis capabilities focusing on particle data have been developed (LANL/ANL collaboration, J. Ahrens, J. Woodring, P. Fasel, K. Heitmann, S. Habib, J. Insley...)
  - Parallel reader, halo finder, sub-sampler, embedded plotting and analysis routines
- **yt (open source)**
  - Focus on astrophysical data, AMR, particles also supported, started by Enzo community (SLAC), expanding fast
  - Parallel halo finder, merger trees, ...
  - Lives also within ParaView now



Some Recent Science Highlights

- **Collaboration members have very successful track record in using N-body simulations to explore new physics and analyze survey data**
- **Cross-laboratory collaborations are established, more being planned**
- **World-leading capabilities exist**

- **Galaxy Surveys**
  - **The clustering of massive galaxies at z~0.5 from the first semester of BOSS data**, M. White et al. ApJ 728, 126 (2011), **LBNL**
  - **Cosmological constraints from galaxy clustering and the the mass-to-number ratio of galaxy clusters,** J. Tinker, E. Sheldon, R. Wechsler et al., ApJ (2011), **BNL/SLAC**
  - **Statistics of satellite galaxies around Milky Way-like hosts,** M. Busha, R. Wechsler et al, ApJ (2011), **SLAC**
  - **The real-space clustering of luminous red galaxies around z<0.6 quasars in the Sloan Digital Sky Survey**, N. Padmanabhan, M. White, et al., MNRAS 397, 1862 (2009), **LBNL**
- **Weak Lensing**
  - **Probing cosmology with weak lensing peak counts,** J. Kratochvil, Z. Haiman, M. May, Phys. Rev. D81, 043519 (2010), **BNL**
  - **Cosmological information in weak lensing peaks,** X. Yang, J. Kratochvil, S. Wang, E. A. Lim, Z. Haiman, M. May, Phys. Rev. D84, 043529 (2011), **BNL**

U.S. DEPARTMENT OF ENERGY | Office of Science

Some Recent Science Highlights

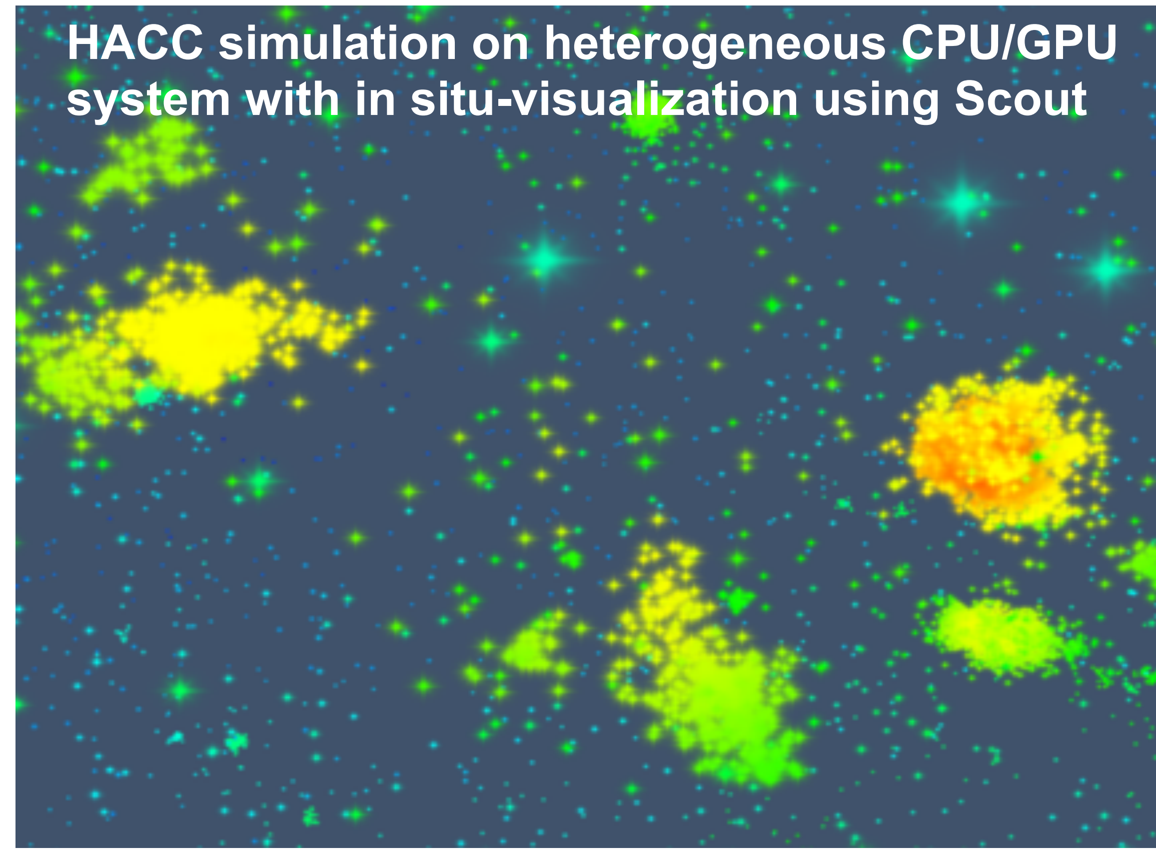
- **Baryon Acoustic Oscillations**
  - **Constraining anisotropic baryon oscillations**, N. Padmanabhan, M. White, Phys. Rev. D77, 123540 (2008), **LBNL**
  - **Non-linear structure formation and the acoustic scale,** H.-J.Seo, E.Siegal, D.Eisenstein, M.White, ApJ 686, 13 (2008), **LBNL**
  - **Calibrating the baryon oscillation ruler for matter and halos,** N. Padmanabhan, M. White, Phys. Rev. D80, 063508 (2009), **LBNL**
- **Baryon Acoustic Oscillations in the Lyman alpha Forest**
  - **The acoustic peak in the Lyman alpha forest,** A. Slosar, S. Ho, M. White, T. Louis, JCAP 10, 019 (2009), **BNL/LBNL**
  - **Particle mesh simulations of the Lyman-alpha forest and the signature of baryon acoustic oscillations in the intergalactic medium,** M. White, A. Pope, J. Carlson, K. Heitmann, S. Habib, P. Fasel, D. Daniel, Z. Lukic, ApJ 713, 383 (2010), **LBNL/ANL**
- **Redshift Space**
  - **Testing cosmological structure formation using redshift-space distortions,** W. Percival, M. White, MNRAS 393, 297 (2009), **LBNL**
  - **Towards an accurate model of the redshift space clustering of halos in the quasilinear regime,** B. Reid, M. White, MNRAS, in press, **LBNL**

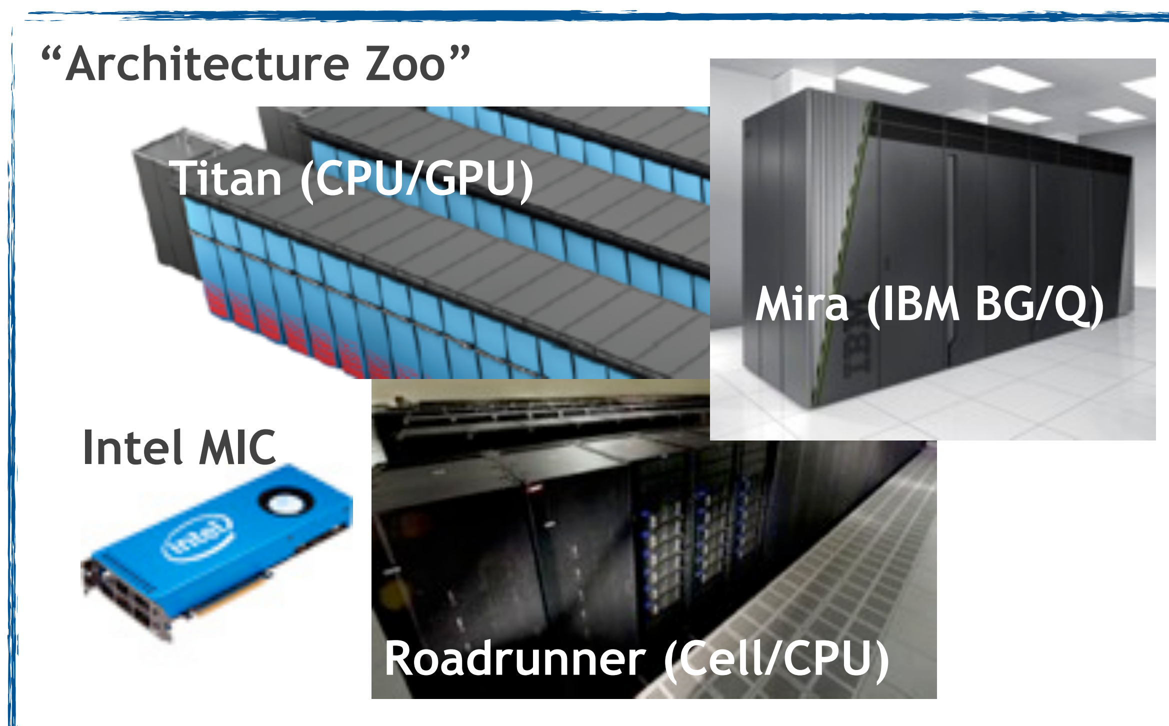
Some Recent Science Highlights

- **Clusters and Groups**
  - **Mass function predictions beyond LCDM,** S. Bhattacharya, K. Heitmann, M. White, Z. Lukic, C. Wagner, S. Habib, ApJ 732, 122 (2011), **ANL/LBNL**
  - **Cosmological constraints from the SDSS maxBCG cluster catalog,** E. Rozo, R. Wechsler et al. ApJ, 708, 645 (2010), **SLAC**
  - **The structure of halos: Implications for group and cluster cosmology,** Z. Lukic, D. Reed, S. Habib, K. Heitmann, ApJ 692, 217 (2009), **ANL**
- **Theory**
  - **The Coyote Universe III: Simulation suite and precision emulator for the nonlinear matter power spectrum,** E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, B. Williams, ApJ 713, 1322 (2010), **ANL/LBNL**
  - **The Coyote Universe I: Precision determination of the nonlinear matter power spectrum,** K. Heitmann, M. White, C. Wagner, S. Habib, D.Higdon, ApJ 713, 1322 (2010), **ANL/LBNL**
  - **The Cosmic Code Coparison Project,** K. Heitmann, Z. Lukic, P. Fasel, S. Habib, M. Warren, M. White, J. Ahrens, L. Ankeny, R. Armstrong, B. O'Shea, P. Ricker, V. Springel, J. Stadel, H. Trac, Compt. Sci. Dis. 1, 015003 (2008), **ANL/LBNL**

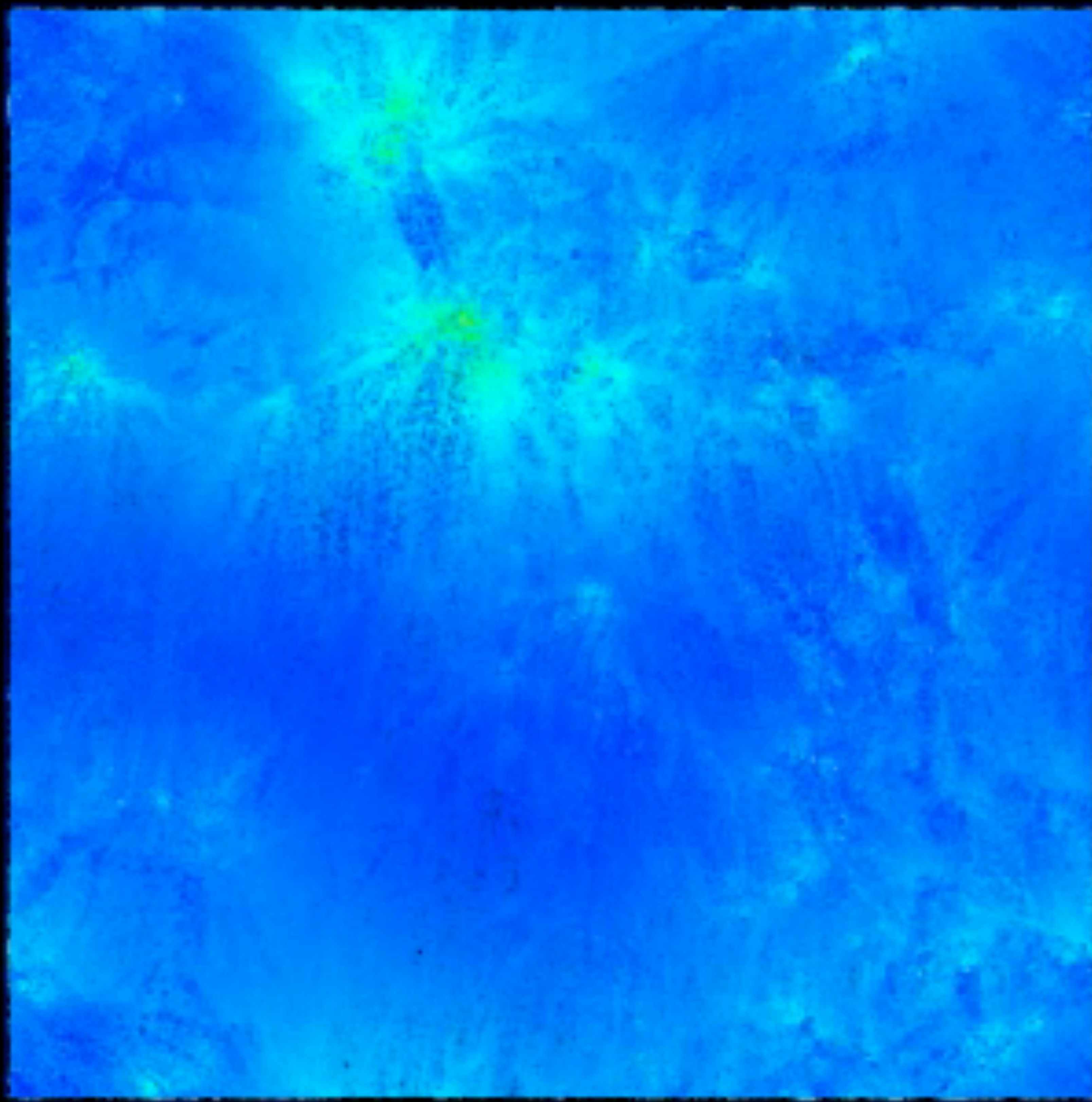
Challenges for the Future

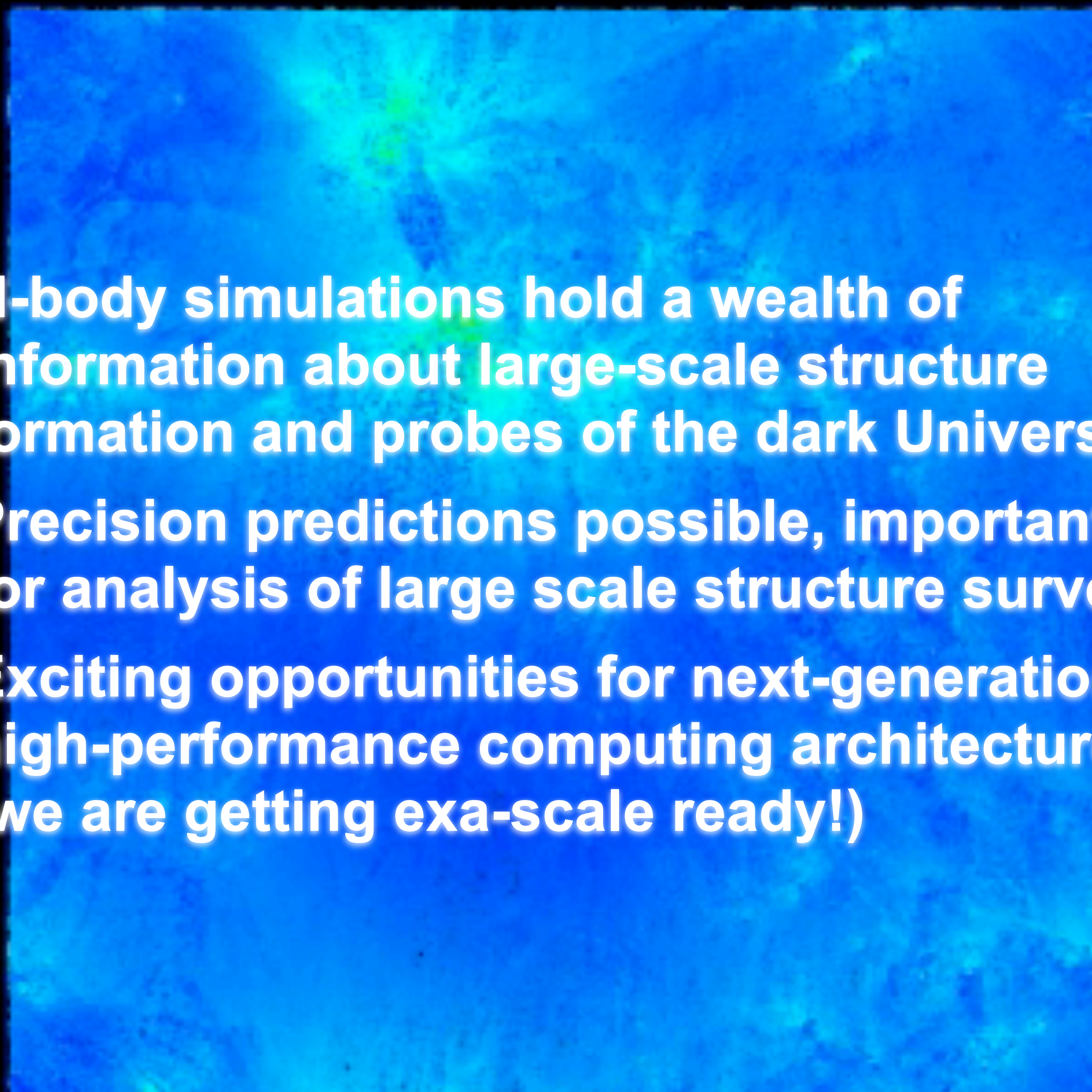
- Large cosmological volumes with ~ trillion particles are needed for ongoing and future large scale surveys
- Hundreds to thousands of realizations, not only for one cosmology!
- Coverage of cosmological model space
- Error control in simulations
- Supercomputer architectures are changing! Codes need to be flexible, close collaboration with computer science community very beneficial
- Simulation analysis, close collaboration with CS community very beneficial
- Data serving, close collaboration with CS community very beneficial
- Outstanding opportunity for HEP and ASCR to work close together and make an impact!
- National Laboratories poised to carry out an effort that can benefit the community





U.S. DEPARTMENT OF ENERGY | Office of Science





- N-body simulations hold a wealth of information about large-scale structure formation and probes of the dark Universe
- Precision predictions possible, important for analysis of large scale structure surveys
- Exciting opportunities for next-generation high-performance computing architectures (we are getting exa-scale ready!)